JVI Journals.ASM.org

Encapsidation of Host RNAs by Cucumber Necrosis Virus Coat Protein during both Agroinfiltration and Infection

Kankana Ghoshal,[a] Jane Theilmann,[b] Ron Reade,[b] Ajay Maghodia,[b] D'Ann Rochon[a,b]

University of British Columbia, Faculty of Land and Food Systems, Vancouver, British Columbia, Canada[a]; Agriculture and Agri-Food Canada, Pacific Agri-Food Research Centre, Summerland, British Columbia, Canada[b]

ABSTRACT

Next-generation sequence analysis of virus-like particles (VLPs) produced during agroinfiltration of cucumber necrosis virus (CNV) coat protein (CP) and of authentic CNV virions was conducted to assess if host RNAs can be encapsidated by CNV CP. VLPs containing host RNAs were found to be produced during agroinfiltration, accumulating to approximately 1/60 the level that CNV virions accumulated during infection. VLPs contained a variety of host RNA species, including the major rRNAs as well as cytoplasmic, chloroplast, and mitochondrial mRNAs. The most predominant host RNA species encapsidated in VLPs were chloroplast encoded, consistent with the efficient targeting of CNV CP to chloroplasts during agroinfiltration. Interestingly, droplet digital PCR analysis showed that the CNV CP mRNA expressed during agroinfiltration was the most efficiently encapsidated mRNA, suggesting that the CNV CP open reading frame may contain a high-affinity site or sites for CP binding and thus contribute to the specificity of CNV RNA encapsidation. Approximately 0.09% to 0.7% of the RNA derived from authentic CNV virions contained host RNA, with chloroplast RNA again being the most prominent species. This is consistent with our previous finding that a small proportion of CNV CP enters chloroplasts during the infection process and highlights the possibility that chloroplast targeting is a significant aspect of CNV infection. Remarkably, 6 to 8 of the top 10 most efficiently encapsidated nucleus-encoded RNAs in CNV virions correspond to retrotransposon or retrotransposon-like RNA sequences. Thus, CNV could potentially serve as a vehicle for horizontal transmission of retrotransposons to new hosts and thereby significantly influence genome evolution.

IMPORTANCE

Viruses predominantly encapsidate their own virus-related RNA species due to the possession of specific sequences and/or structures on viral RNA which serve as high-affinity binding sites for the coat protein. In this study, we show, using next-generation sequence analysis, that CNV also encapsidates host RNA species, which account for ~0.1% of the RNA packaged in CNV particles. The encapsidated host RNAs predominantly include chloroplast RNAs, reinforcing previous observations that CNV CP enters chloroplasts during infection. Remarkably, the most abundantly encapsidated cytoplasmic mRNAs consisted of retrotransposon-like RNA sequences, similar to findings recently reported for flock house virus (A. Routh, T. Domitrovic, and J. E. Johnson, Proc Natl Acad Sci U S A 109:1907–1912, 2012). Encapsidation of retrotransposon sequences may contribute to their horizontal transmission should CNV virions carrying retrotransposons infect a new host. Such an event could lead to large-scale genomic changes in a naive plant host, thus facilitating host evolutionary novelty.

Viruses generally selectively package their own genomic RNA over host RNA. In many cases, virus capsid proteins have been found to bind specific high-affinity sequences and/or structures on viral RNA, which helps to preclude encapsidation of host RNAs (1–17). In the case of the satellite tobacco necrosis virus (STNV) and MS2 bacteriophage, specific RNA sequences may not exist; rather, the genome may contain numerous distinct but functionally similar packaging sites (e.g., small stem-loops) that are bound simultaneously but independently by multiple coat protein (CP) subunits (4, 18, 19). In addition, it has been suggested for some viruses that assembly occurs within viral factories consisting predominantly of viral RNA, thus making encapsidation of host RNA less likely. For example, the replicase-associated 2C ATPase encoded by the poliovirus genome interacts with the capsid protein, and this leads to morphogenesis at the replication site (20). Many viruses are known to encapsidate host RNAs when the coat protein is expressed under artificial conditions. One such example is MS2 phage, which encapsidates a 16S rRNA precursor during expression from a plasmid in bacterial cells (21). Pariacoto virus (PaV) as well as flock house virus (FHV), both of which are nodaviruses, can produce virus-like particles (VLPs) during artificial expression in insect cells (22, 23). Potato leaf roll virus also forms VLPs when the coat protein is expressed in insect cells and encapsidates host RNAs in the size range of 80 to 200 nucleotides (nt) (24). Transient expression of brome mosaic virus (BMV) CP in the absence of replicase also results in encapsidation of host RNA species (25). Siegel (26) was the first to show that a plant virus,

Received 5 June 2015 **Accepted** 3 August 2015

Accepted manuscript posted online 12 August 2015

Citation Ghoshal K, Theilmann J, Reade R, Maghodia A, Rochon D. 2015. Encapsidation of host RNAs by cucumber necrosis virus coat protein during both agroinfiltration and infection. J Virol 89:10748–10761. doi:10.1128/JVI.01466-15.

Editor: A. Simon

Address correspondence to D'Ann Rochon, dann.rochon@agr.gc.ca.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/JVI.01466-15.

tobacco mosaic virus (TMV), encapsidates host RNA during infection, with approximately 2 to 2.5% of total virion RNA of the U2 strain of TMV corresponding to host RNA sequences and predominantly chloroplast RNA. Later, Rochon and Siegel (15) showed that chloroplast DNA transcripts, excluding rRNAs, are encapsidated by the U2 strain of TMV. FHV has recently been found to encapsidate host RNAs during infection, with up to 1% of total virion RNA corresponding to host RNA sequences (23). Thus, although mechanisms are in place to ensure encapsidation of predominantly viral RNA during infection, in a few cases, host RNAs have also been shown to be encapsidated, albeit at a significantly lower level.

Cucumber necrosis virus (CNV) is a member of the *Tombusvirus* genus in the family *Tombusviridae* (27–29). As such, the genome is positive-sense, single-stranded RNA encapsidated by a T=3 icosahedral shell. The genome is approximately 4.7 kb, and the 34-nm capsid is comprised of 180 identical copies of a single 41-kDa CP. Besides the CP, the CNV genome encodes 4 additional proteins (Fig. 1A): an auxiliary replicase factor (p33), an RNA-dependent RNA polymerase (p92), the cell-to-cell-movement protein (p21), and a silencing suppressor (p20). The CNV CP is translated from a subgenomic RNA (sgRNA1) during infection, as are p21 and p20 (sgRNA2). p33 is translated from full-length genomic RNA, and p92 is a readthrough product of the leaky amber termination codon of the p33 open reading frame (ORF) (29–33).

An X-ray crystal structure of the CNV particle has been determined at a 2.9-Å resolution (34), and the particle structure has also been determined by cryo-electron microscopy (cryo-EM) at a 12-Å resolution (35). The cryo-EM structure shows that the capsid consists of both an inner shell and an outer shell, with the inner shell forming a T=1-like scaffold that may be used for particle assembly. The structures of the particle and CP of CNV are similar to those of the tombusvirus tomato bushy stunt virus (TBSV) (35). The CP subunit of CNV consists of 3 major domains: an inward-facing RNA binding domain (R); the shell domain (S), which forms the outermost shell of the particle; and the P domain, which projects from the particle surface (Fig. 1A). A flexible 34-amino-acid (aa) arm connects the R domain to the S domain, and a 5-aa hinge tethers the S domain to the P domain. Deletion analysis of the R domain has shown that specific R-domain sequences are required for T=3 icosahedral symmetry, as T=1 particles are formed in their absence (35–37). Additionally, it has been shown that a KGKKGK sequence at the C-terminal region of the R domain is required for encapsidation of full-length CNV RNA, at least in part due to its role in the neutralization of negatively charged phosphates of encapsidated RNA (37).

Approximately 1 to 5% of the CNV CP in a CNV-infected cell is present in chloroplasts (38). Although the biological significance of chloroplast targeting is not known, it was suggested that CNV particles may uncoat on the cytoplasmic side of the chloroplast outer membrane as CP subunits are reeled into the chloroplast during CP uptake (38). Using green fluorescent protein (GFP) fusion protein constructs, it has been determined that most of the arm domain plus 4 aa of the S domain are minimally required for chloroplast targeting (38). Coat protein-GFP fusion protein constructs have also been shown to target mitochondria, wherein 38 aa of the N-terminal region of the CP are sufficient for mitochondrial targeting (39). We have previously shown that 3 specific cleavage events occur during chloroplast targeting of the

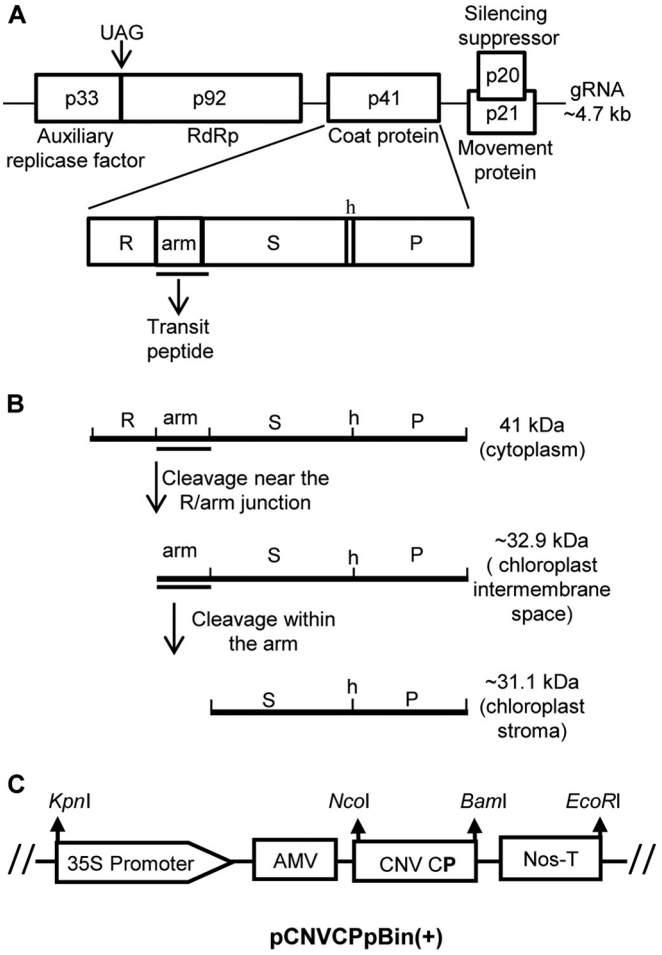

FIG 1 Features of the CNV genome, CP cleavage products, and CNV CP expression vector used in this study. (A) Genome organization of CNV showing the 5 open reading frames and the proteins that they encode. The CP ORF is expanded, and the 3 main domains (R, S, and P) along with the arm region and hinge (h) are shown. The approximate region of the CP that contains the chloroplast transit peptide-like sequence is underlined. RdRp, RNA-dependent RNA polymerase; gRNA, genomic RNA. (B) Diagrammatic representation of 3 CP cleavage events that take place during targeting of the CNV CP to chloroplasts (47, 48). The first cleavage is near the R/arm junction and results in an N-terminal ~32.9-kDa truncated CP where the chloroplast transit peptide is now at the N terminus of the protein. This cleaved species resides in the chloroplast intermembrane space or mitochondrial matrix. The second cleavage is near the arm/S junction and occurs within the stroma (where this 31.1-kDa species resides) likely via the stromal processing peptidase. (C) Diagrammatic representation of the structure of pCNVCPpBin(+). The 35S promoter, the alfalfa mosaic virus (AMV) translational enhancer, and the nopaline synthase transcription termination signal (Nos-T) are shown along with the position of the CNV CP. The restriction enzyme cleavage sites used for cloning purposes are also indicated.

CNV CP (38, 39) (Fig. 1B). Cleavage near the R/arm boundary, which may occur in the cytoplasm or mitochondria, gives rise to a 32.9-kDa protein. This protein is functionally equivalent to a chloroplast preprotein, in that it contains an N-terminal chloroplast transit peptide-like sequence of at least 38 aa and is capable of targeting chloroplasts. A second cleavage was found to occur near the arm/shell junction during chloroplast uptake into the stroma, giving rise to a protein of 31.1 kDa. A third minor cleavage event (not shown in Fig. 1) occurs in the mitochondria in the R domain approximately 25 to 35 aa from the N terminus (38, 39).

During studies to analyze the biological significance of chloroplast targeting of the CNV CP and to analyze the viral origin of an assembly sequence(s), we have found that *Nicotiana benthamiana* plants agroinfiltrated with CNV CP accumulate both T=3- and T=1-like particles that encapsidate specific host RNA species, here referred to as VLPs. Using next-generation sequencing (NGS) analysis of RNA extracted from RNase-treated VLPs, we found that the majority of the encapsidated host RNA is of chloroplast origin, but host RNAs of mitochondrial and nuclear origin are also encapsidated. rRNAs from each of these organellar compartments were included among the host RNAs found to be encapsidated in VLPs. The CNV CP mRNA generated during agroinfiltration was additionally encapsidated and, compared with other efficiently encapsidated host RNAs, had the highest relative encapsidation efficiency. Additionally, the TBSV p19 mRNA expressed during coagroinfiltration with CNV CP was also encapsidated relatively efficiently, suggesting that the CNV CP ORF and the CNV p20 ORF (CNV p20 is the homolog of TBSV p19) each contains a high-affinity CNV CP binding site used for encapsidation. Additionally, we have found that highly purified, RNase-treated CNV particles extracted from infected leaves also contain host RNA (~0.1% of total virion RNA), again, with the majority of the host RNA corresponding to chloroplast RNA. As has been shown for FHV (23), retrotransposon RNA sequences are also encapsidated and represent approximately 0.5 to 1.3% of total encapsidated host RNAs and, remarkably, the highest proportion of encapsidated nonribosomal nucleus-encoded RNA. These results suggest a means whereby retrotransposons can be horizontally transferred to other CNV hosts and, therefore, a means by which viruses can significantly influence the evolution of host genomes.

MATERIALS AND METHODS

Plasmid construction. The pCNVCPpBin(+) construct containing the CNV coat protein ORF in the *Agrobacterium tumefaciens* binary plasmid vector pBin(+) was prepared in two steps. In the first step, the CNV CP ORF was cloned into the intermediate vector pBBI525 between the dual 35S promoter and the NOS terminator to form pBBI525/CNVCP. The CNV CP ORF was PCR amplified by using the forward primer CNV-040 and the reverse primer CNV-279 (see Table S1 in the supplemental material) from a previously described full-length infectious CNV cDNA clone, pK2/M5 (30). The PCR product was inserted in the NcoI and BamHI sites of pBBI525 by bimolecular ligation. The 35S promoter, the CNV CP ORF, and the NOS terminator from pBBI525 were digested with KpnI and EcoRI and inserted into similarly digested pBin(+). The final structure of the pCNVCPpBin(+) construct is shown in Fig. 1C.

***Agrobacterium*-mediated transient expression.** *Agrobacterium*-mediated transient expression of pCNVCPpBin(+) was performed as described previously (40).

Virus and VLP purification. Wild-type (WT) CNV was propagated in *N. benthamiana* plants mechanically inoculated with infectious transcripts prepared from pK2/M5 as described previously (30). Infected leaves from transcript-inoculated plants were used to inoculate a large batch of 4- to 6-week-old *N. benthamiana* plants. VLPs were produced by agroinfiltration of *N. benthamiana* leaves with pCNVCPpBin(+). The TBSV p19 protein was coexpressed in the agroinfiltrated plant cells to suppress induction of gene silencing (41), thereby allowing optimal expression of pCNVCPpBin(+). CNV virions and pCNVCPpBin(+) VLPs were extracted using differential centrifugation as previously described (35). WT CNV particles were electrophoresed through 1% (wt/vol) agarose gels in TB buffer (45 mM Tris, 45 mM borate, pH 8.3) as described previously (42). Virions were stained with ethidium bromide (EtBr) in the presence of TB buffer containing 1 mM EDTA and photographed under UV illumination. Electrophoresis of VLPs from CNV CP-infiltrated plants was conducted using 2% (wt/vol) agarose gels in TB buffer for 1.5 h (36). Known concentrations of WT CNV particles (as determined by spectrophotometry) were used as mass standards. The yield of VLPs was typically approximately 1/60 of the yield of CNV virions.

For the smaller-scale purification of CNV and VLPs, a procedure that involved polyethylene glycol precipitation and only low-speed centrifugation was utilized (36, 43).

RNase treatment of CNV virions and VLPs. CNV virions and VLPs were treated with RNase A (Invitrogen), when indicated, at a concentration of 0.02 ng/μl for 30 min at room temperature. Prior to RNA extraction using phenol chloroform and SDS in the presence of EDTA as previously described (30), the mixtures were adjusted to 20 mM β-mercaptoethanol to inhibit RNase activity during the purification procedure. Pilot experiments showed that free RNA was readily digested under these conditions, whereas virion RNA remained intact. Purified RNA was then treated with DNase I (Qiagen) as described below for total leaf RNA.

Transmission electron microscopy of CNV virions and VLPs. Transmission electron microscopy of uranyl acetate-stained virions was conducted as described previously (36).

SDS-PAGE and Western blot analysis. Total leaf protein, VLPs, and virions were subjected to SDS-PAGE and Western blot analysis as previously described (40) using a polyclonal antibody raised against the bacterially expressed S and P domains of the CNV CP (antibody SP).

Total leaf RNA purification, electrophoresis, and Northern blot analysis. Total leaf RNA was extracted using an RNeasy plant kit (Qiagen) following the manufacturer's protocol, including DNase I treatment. RNA was electrophoresed through 1% agarose gels in 0.5× TBE (44.5 mM Tris, 44.5 mM boric acid, 1 mM EDTA, pH 8) and visualized by staining with EtBr as previously described (42). For Northern blot analysis, virion RNA was separated in a 1% formaldehyde-agarose gel, transferred to Zeta-probe GT blotting membranes (Bio-Rad), and hybridized with a randomly primed ^{32}P-labeled cDNA probe specific for uninfected total leaf RNA using a Random Primers DNA labeling system (Invitrogen).

N-terminal peptide sequencing. The N termini of the pCNVCPpBin(+) 41-kDa, ~33-kDa, and ~31-kDa proteins present were determined by Edman degradation analyses by the Protein Facility of the Iowa State University Office of Biotechnology using a PerkinElmer Applied Biosystems model 494 Procise protein/peptide sequencer. Protein species were obtained from pCNVCPpBin(+)-infiltrated leaf tissue by immunoprecipitation using a CNV CP-specific polyclonal antibody. The proteins were electrophoresed through an SDS-polyacrylamide gel, and individual CP species were excised from the gel prior to Edman degradation analysis.

NGS of RNA. Two different CNV virion RNA preparations or VLP RNA prepared as described above was subjected to NGS using the Ion Torrent semiconductor or Illumina sequencing platform. For Ion Torrent sequencing (Ion Proton; Life Technologies), reads were single end and the length between fragments varied by up to 200 bases. The library preparation method utilized a total RNA sequencing kit (v2; Life Technologies) with no poly(A) enrichment. For Illumina sequencing, the length was 75 nucleotides. The library preparation method utilized a TruSeq Stranded mRNA LT kit (Illumina) with no poly(A) enrichment. CNV virions were obtained from 2 independent virus purifications from different batches of infected leaf tissue. In the first case, virions were not treated with RNase and total RNA was sequenced using Ion Torrent sequencing. In the second case, virions were treated with RNase prior to RNA extraction, and sequencing was by use of the Illumina technology. CNV VLPs were treated with RNase prior to RNA extraction, and sequencing was performed by use of the Ion Torrent technology. Sequencing was outsourced to Applied Biological Materials, Inc. (Richmond, BC, Canada). Prior to sequencing, the quality of the RNA was determined using an Agilent 2100 bioanalyzer (Agilent Technologies, Inc.). Poor-quality reads containing any base with a Phred score of <20 and reads containing less than 50 bases were re-

moved. Three data sets corresponding, respectively, to CNV VLP RNA and to untreated and RNase-treated CNV virions were obtained. These data sets are referred to as CNVVLP-RNAseq, CNV-RNAseq1, and CNV-RNAseq2, respectively.

Alignment of reads from next-generation sequencing. Alignment of the reads was conducted using the CLC Genomics Workbench (v6.5 or v7.5). A similarity fraction of 0.95 and a length fraction of 0.5 were used for all mappings except for the rRNAs, where the similarity fraction was 0.95 and the length fraction was 0.95. The mismatch cost was set to 2, the insertion cost was set to 3, and the deletion cost was set to 3. The NCBI GenBank accession numbers of the genes or genomes used for alignment of the reads were M25270 for the full-length sequence of CNV RNA (29), NC_001554.1 for the TBSV p19 sequence (44), Z00044 for the 23S and 16S rRNA genes derived from the complete *Nicotiana tabacum* chloroplast genome (45), and BA000042 for the 26S and 18S rRNA genes derived from the *N. tabacum* mitochondrial genome (46). GenBank accessions AF479172 and AJ236016 were used for the complete *N. tabacum* 26S and 18S cytoplasmic rRNAs, respectively. The *N. benthamiana* transcriptome (v5) was downloaded from http://sydney.edu.au/science/molecular_bioscience/sites/benthamiana/ (47). Sequences that mapped to the *N. benthamiana* transcriptome were identified on the basis of the annotations in the downloaded *N. benthamiana* transcriptome. The accuracy of several of the annotations was confirmed by BLAST analysis against the sequences in the NCBI database and adjusted accordingly (see Fig. S1 to S3 in the supplemental material).

RT-PCR analysis. One hundred nanograms of mock total leaf RNA, 100 ng of pCNVCPpBin(+) VLP RNA, or 5 μg of WT CNV RNA obtained as described above was reverse transcribed into cDNA using SuperScript III reverse transcriptase (RT; Invitrogen) and gene-specific reverse primers (see Table S1 in the supplemental material). cDNA was amplified using *Taq* polymerase (Invitrogen) and the gene-specific primers indicated in Table S1 in the supplemental material using standard PCR protocols. RT-PCR products were analyzed by electrophoresis through 1% agarose gels buffered in TAE (Tris-acetate-EDTA) followed by staining with ethidium bromide.

ddPCR analysis. SuperScript III RT (Invitrogen) and gene-specific reverse primers for cDNA synthesis (see Table S1 in the supplemental material) were used for reverse transcription. Serial dilutions of the cDNAs were prepared, and droplet digital PCR (ddPCR) was performed using a QX200 droplet digital PCR system (Bio-Rad) according to the manufacturer's protocol.

Calculation of REE. Calculation of the encapsidation efficiency (REE) of CNV CP mRNA relative to that of other encapsidated pCNVCPpBin (+) VLP RNAs was conducted as described previously (10) using the results of ddPCR. The ratio of the number of molecules of the indicated RNA in virion RNA with respect to that of CNV CP mRNA was divided by the ratio of the number of molecules of the indicated RNA in total leaf RNA with respect to that of CNV CP mRNA. The result was multiplied by 100 and is shown as a percentage. A similar formula was applied for calculating the REE of host RNAs.

RESULTS AND DISCUSSION

Agroinfiltrated CNV CP encapsidates host RNA species. CNV CP was cloned into the *A. tumefaciens* binary vector pBin(+) and used to agroinfiltrate *N. benthamiana* leaves. This construct, pCNVCPpBin(+) (Fig. 1C), was coagroinfiltrated with the tomato bushy stunt virus (TBSV) p19 suppressor of silencing (41) to suppress silencing of CNV CP mRNA and, consequently, increase the levels of CNV CP. At 5 days postinfection (dpi), leaves were collected and subjected to virus purification utilizing a method previously used for the purification of CNV particles from infected plants (36). Agarose gel electrophoresis followed by ethidium bromide staining of purified material showed the presence of two distinct virus-like particles (VLPs) (Fig. 2A): one species comigrated with authentic CNV T=3 particles, whereas the other species migrated at a higher rate similar to that of the T=1 particles formed by CNV mutants lacking R-domain sequences (36). The putative VLPs were further analyzed by negative staining followed by electron microscopy (Fig. 2B). It can be seen that particles similar in size to authentic 34-nm CNV T=3 particles were evident, along with particles of a smaller diameter of approximately 23 nm, similar in size to previously described CNV T=1 particles (35, 36). In addition, particles with apparent defects in structure, such as the one indicated in Fig. 2B, panel ii, where the particle appears to contain a hole with missing CP subunits, were also observed.

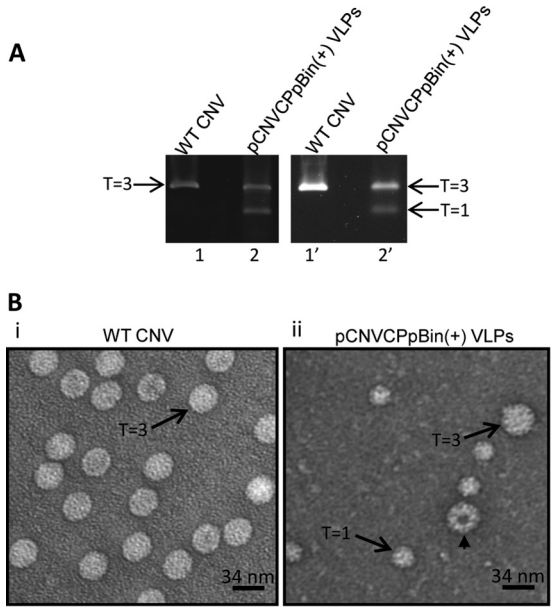

FIG 2 Features of virus-like particles extracted from pCNVCPpBin(+)-agroinfiltrated plants. (A) Agarose gel electrophoresis of virus-like particles extracted from pCNVCPpBin(+)-agroinfiltrated plants. Particles were extracted from CNV-infected leaves (lanes 1 and 1′) or from pCNVCPpBin(+)-infiltrated leaves (lanes 2 and 2′). CNV particles and the VLPs extracted from pCNVCPpBin(+)-infiltrated leaves were electrophoresed through separate 2% agarose gels and either stained with EtBr in the presence of 1 mM EDTA (left) or stained with Sypro Ruby (right). The positions of the T=3 and T=1 particles are shown. The yield of VLPs was typically approximately 1/60 of the yield of CNV virions. (B) Transmission electron microscopy analysis of VLPs. WT CNV particles (i) and particles extracted from pCNVCPpBin(+)-infiltrated plants (ii) are shown. Arrows, the presence of T=3 particles of WT CNV and T=3-like and T=1-like particles extracted from pCNVCPpBin(+)-infiltrated plants; arrowhead, a T=3-like particle with an apparent structural defect.

pCNVCPpBin(+) VLPs were subjected to RNase treatment to eliminate any RNA that may have copurified with the virus particles, and then RNA was extracted from the treated VLPs and analyzed by agarose gel electrophoresis. Four discrete species of approximately 3.3 kb, 2.0 kb, 1.8 kb, and 1.4 kb were observed along with a mixture of smaller species ranging in size from ~0.5 to 1.3 kb (Fig. 3A, lane 1). The pattern of RNA species observed did not differ between RNase-treated and untreated VLPs (data not shown). The approximately 3.3-kb species comigrated with host cytoplasmic 26S rRNA, and the ~1.8-kb species comigrated with host cytoplasmic 18S rRNA. The ~2.0-kb RNA, the 1.4-kb RNA, and the series of smaller species did not have a clear counterpart in

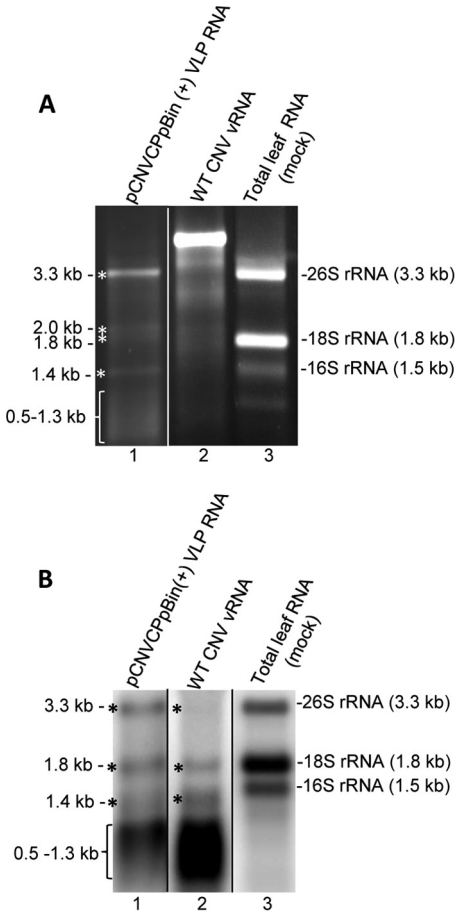

FIG 3 Agarose gel electrophoresis and Northern blot analysis of host RNA species present in CNVCPpBin(+) VLPs and CNV particles. (A) Agarose gel electrophoresis of RNA extracted from RNase-treated pCNVCPpBin(+) particles. Lane 1, RNA extracted from pCNVCPpBin(+) VLPs; lane 2, CNV virion RNA (vRNA); lane 3, total RNA from uninfected leaves (mock infected). The sizes of cytoplasmic 26S rRNA (3.3 kb), 18S rRNA (1.8 kb), and 16S rRNA (1.5 kb) are indicated on the right. Asterisks, the major RNAs of a discrete size extracted from pCNVCPpBin(+) VLPs; brackets, a series of smaller RNAs ranging from ~0.5 to ~1.3 kb. (B) Northern blot analysis. Fifty nanograms of RNA extracted from RNase-treated CNVCPpBin(+) VLPs (lane 1) or 2.5 μg of RNA extracted from RNase-treated CNV particles was electrophoresed through a 1% denaturing agarose gel and blotted onto a Zeta-probe membrane (Bio-Rad). The probe was a ^{32}P-labeled randomly primed cDNA to total leaf RNA extracted from *N. benthamiana*. Lane 3, 50 ng of total leaf RNA used as a control. Lanes 1 and 3 were exposed for approximately 1/12 the time that lane 2 was exposed.

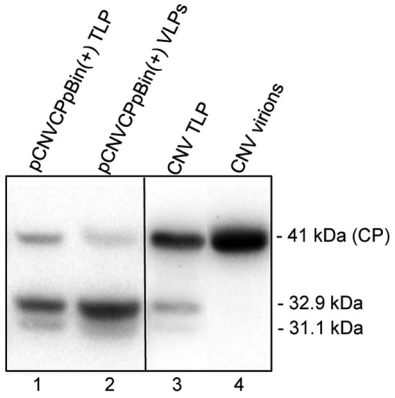

FIG 4 Western blot analysis of CNVCPpBin(+) total leaf protein and VLP protein. *N. benthamiana* leaves were agroinfiltrated with pCNVCPpBin(+), and 5 days following agroinfiltration, total leaf protein (lane 1) or purified VLPs (lane 2) were denatured and electrophoresed through a NuPAGE gel and subjected to Western blot analysis using a polyclonal antibody specific to the CNV CP shell and protruding domains (antibody SP). Lane 3, total leaf protein extracted from CNV-infected leaves; lane 4, CNV virion protein used as a size control (see the text). The numbers on the right correspond to the sizes of full-length CNV CP (41 kDa), the 32.9-kDa CP cleavage product associated with the chloroplast intermembrane space and with mitochondria, and the 31.1-kDa cleavage product associated with the chloroplast stroma (47, 48). The exact sizes of the cleavage products were determined by Edman degradation analysis of protein immunoprecipitated from pCNVCPpBin(+)-agroinfiltrated leaves (see the text).

total uninfected leaf RNA, suggesting that these species may be selectively encapsidated during infection. To further analyze the host origin of pCNVCPpBin(+) VLP RNA species, a randomly primed cDNA probe for total uninfected *N. benthamiana* leaf RNA was hybridized to VLP RNA, which was electrophoresed through a denaturing agarose gel and blotted onto a Zeta-probe GT membrane (Bio-Rad) (Fig. 3B). The results show that the 3.3-kb, 1.8-kb, and 1.4-kb species visible on ethidium bromide-stained gels each hybridized to host RNA. On the basis of the sizes of the RNAs, the 3.3-kb species may be 26S rRNA and the 1.8-kb species may be 18S rRNA, but this requires further confirmation. The putative identities of the remaining 2.0-kb species, the 1.4-kb species, and the series of smaller RNA species which do not have detectable counterparts in total leaf RNA are unknown. Collectively, the results indicate that pCNVCPpBin(+) VLP RNA corresponds to host RNA encapsidated by pCNVCPpBin(+) CP following agroinfiltration. Thus, although it is generally believed that viral CPs selectively encapsidate viral RNA, agroinfiltrated CNV CP efficiently encapsidates host RNAs, suggesting as described further below that encapsidation specificity may occur preferentially during the infection process.

To eliminate the unlikely possibility that uninfiltrated *N. benthamiana* harbors VLPs, uninfiltrated leaves were subjected to virus purification using the same amount of tissue and the methods used for purification of the pCNVCPpBin(+) VLPs. No virus particles were observed, as determined by agarose gel electrophoresis (data not shown). As well, purified material was subjected to the virion RNA extraction procedure utilized for purification of CNV RNA and pCNVCPpBin(+) VLP RNA, and no RNA was observed following agarose gel electrophoresis (data not shown).

pCNVCPpBin(+) VLPs are comprised of full-length CNV CP as well as 2 CP cleavage products associated with chloroplasts or mitochondria. Total protein from pCNVCPpBin(+) VLP preparations was analyzed by SDS-PAGE followed by Western blotting analyses to determine the sizes of the CP species present in pCNVCPpBin(+) VLPs (Fig. 4). In addition to the full-length 41-kDa CNV CP species, 2 additional species of approximately 32.9 kDa and 31.1 kDa were observed. The full-length 41-kDa species as well as the 2 cleavage products associated with VLPs comigrated with proteins detected in total leaf protein from pCNVCPpBin(+)-agroinfiltrated plants (Fig. 4, compare lanes 2 and 1, respectively). These 2 cleavage products, in turn, comigrated with the 2 CNV cleavage products that have previously been shown to be associated with chloroplasts or mitochondria (38, 39) and can be detected in total leaf extracts of CNV-

infected plants (Fig. 4, lane 3). As summarized in Fig. 1, the 32.9-kDa species is the result of cleavage near the R/arm junction and has been found to be associated with the chloroplast intermembrane space as well as with mitochondria (38, 39). The 31.1-kDa species is a further cleavage product of the chloroplast 32.9-kDa species and results from cleavage near the arm/S junction in the chloroplast stroma (38, 39). Collectively, the results demonstrate that pCNVCPpBin(+) VLPs are comprised of full-length CNV CP as well as the 2 chloroplast proteins or the mitochondrial cleavage product. Previous studies have shown that CNV T=1 particles can be formed during infection from CNV CP deletion mutants lacking the CP R domain (36, 37), suggesting that it may be possible that the T=1 particles present in pCNVCPpBin(+) VLPs are comprised exclusively of the 32.9-kDa species. Studies have not been conducted to determine if particles lacking both the R and the arm domains can form T=1 particles, so it is not possible to speculate on whether the 31.1-kDa protein can itself form particles. We also note that the 32.9-kDa species might also be present in the T=3-sized particles, as such an observation was previously made with CNV T=3 particles of a mutant that also produced T=1 particles (36).

To precisely identify and map the N termini of the CP cleavage products, pCNVCPpBin(+) CP species were immunoprecipitated with a CNV-specific polyclonal antibody and the two cleavage products were isolated from SDS-polyacrylamide gels and sequenced. Since CNV virions also contain low levels of a similarly sized cleavage product, CNV virions were also used as a source of protein. Edman degradation analysis showed that the N terminus of the 32.9-kDa species corresponds to CNV CP amino acids YAVKG (aa 78 to 82) just downstream from the R/arm junction (yielding a protein with a molecular mass of 32.9 kDa) and the 31.1-kDa product corresponds to CNV CP amino acids SVRIT (aa 93 to 97), corresponding to the first 5 aa of the S domain at the arm/S junction (yielding a protein with a molecular mass of 31.1 kDa).

It is noted that although the 32.9-kDa cleavage product predominates in CNVCPpBin(+) VLPs and in total leaf protein from *N. benthamiana* leaves agroinfiltrated with CNVCPpBin(+), the full-length 41-kDa CNV CP predominates in CNV virions and total leaf protein from infected *N. benthamiana* leaves (Fig. 4). Although we have not directly tested the basis for this, it is likely that full-length CNV CP is readily assembled with virion RNA during infection and this not only stabilizes the 41-kDa species but also precludes chloroplast targeting. However, in pCNVCPpBin(+)-infiltrated plants, full-length CP does not readily assemble into particles and therefore is relatively more available for chloroplast targeting. Our studies have shown that pCNVCPpBin(+)-infiltrated tissue yields, on average, only approximately 5 μg VLP/g leaf tissue, whereas CNV-infected leaves yield approximately 300 μg virions/g leaf tissue.

Identification and characterization of host RNA species encapsidated in pCNVCPpBin(+) VLPs and WT CNV virions. NGS analysis was conducted to determine the identity of host RNA species encapsidated by pCNVCPpBin(+) VLPs and by WT CNV virions during infection. RNA was extracted from RNase-treated VLPs and from WT CNV particles that were either not treated (CNV-nr) or treated (CNV-r) with RNase (18). After quality control analyses, approximately 8.1 million reads from pCNVCPpBin(+) VLP RNA, 9.1 million reads from CNV-nr RNA, and 372 million reads from CNV-r RNA were obtained. The three data sets are referred to as CNVVLP-RNAseq, CNV-RNAseq1, and CNV-RNAseq2, respectively. The reads from CNVVLP-RNAseq were aligned to the *N. benthamiana* transcriptome (v5) from the University of Sydney School of Molecular Bioscience (http://sydney.edu.au/science/molecular_bioscience/sites/benthamiana/). Host sequences were identified on the basis of the annotations in the *N. benthamiana* transcriptome, and several annotations were confirmed or redefined by BLAST analysis using the NCBI database (see Fig. S1 to S3 in the supplemental material). Since both CNV CP and TBSV p19 mRNAs were expressed in agroinfiltrated plants from which pCNVCPpBin(+) VLPs were obtained, the numbers of reads obtained for each sequence were determined using the respective sequences available in the NCBI database (see Materials and Methods). The full-length sequence of CNV RNA was used as a source for the viral genome in the analysis of the CNV-RNAseq1 and CNV-RNAseq2 data. Unmatched reads were then aligned with the *N. benthamiana* transcriptome as described above. To quantify the relative amounts of rRNAs encapsidated, the individual full-length cytoplasmic *N. tabacum* 26S and 18S rRNAs, chloroplast 23S and 16S rRNAs, and mitochondrial 26S and 18S rRNAs were used as reference genes. Figure 5 summarizes the major findings. Of the approximately 8.1 million reads in the CNVVLP-RNAseq data set, approximately 5 million mapped to the *N. benthamiana* transcriptome and 2.9 million remained unassigned. Sequences corresponding to the CNV CP mRNA and the TBSV p19 mRNA accounted for approximately 49,000 (0.93%) and 12,000 (0.24%) of the assigned reads, respectively. These data confirm that *N. benthamiana* RNA species are encapsidated by CNV VLPs and also demonstrate that both TBSV p19 RNA and CNV CP mRNA are encapsidated (Fig. 5A).

Similar analyses using the CNV-RNAseq1 data set showed that of the approximately 9.1 million reads, 8.96 million corresponded to assigned sequences, of which CNV RNA made up 99.29%, and 191,000 were unassigned (Fig. 5B). *N. benthamiana* host RNA sequences accounted for approximately ~63,000 reads (including reads not annotated, i.e., ~9,000), indicating that approximately 0.7% of CNV virion RNA is host derived. For CNV-RNAseq2, approximately 368 million reads were obtained; ~343 million of those corresponded to CNV RNA sequences and approximately 25 million remained unassigned. *N. benthamiana* host RNA sequences accounted for approximately ~296,000 reads (including reads not annotated), indicating that approximately 0.09% of CNV virion RNA is host derived. Since CNV-RNAseq2 data were derived from RNase-treated virions, 0.09% is likely the closer estimate of the contribution of host RNA to total encapsidated virion RNA.

The most abundantly encapsidated RNAs derived from the chloroplast, mitochondrial, and nuclear genomes are summarized in Fig. 6 for the 3 data sets. In the CNV VLPs (Fig. 6A), chloroplast 23S rRNA was the most abundantly encapsidated RNA, with a reads per kilobase (RPK) value of ~761,000. The second most abundantly encapsidated RNA was chloroplast 16S rRNA, with an RPK value of ~417,000. Nuclear and mitochondrial rRNAs were also highly encapsidated, with RPK values ranging from ~44,000 to ~303,000. Ribulose bisphosphate carboxylase large chain (RbcL) mRNA, which is encoded by the chloroplast, had the highest RPK value of the encapsidated mRNAs encoded by the nucleus, chloroplast, and mitochondria. Other chloroplast RNAs also had higher RPK values, with the top 6 chloroplast RNAs having RPK values higher than those of either of the top nucleus-

A CNVVLP-RNAseq

	Reads	% Assigned reads
Total reads	8,154,971	----
Assigned	5,239,655	100%
-CNV CP ORF	48,887	0.93%
-TBSV p19 ORF	12,360	0.24%
-*N. benthamiana*	5,070,108	96.76%
-Not annotated	108,300	2.07%
Unassigned	2,915,316	

B CNV-RNAseq1

	Reads	% Assigned reads
Total reads	9,150,534	----
Assigned	8,959,510	100%
-WT CNV RNA	8,896,001	99.29%
-*N. benthamiana*	54,288	0.61%
-Not annotated	9,221	0.1%
Unassigned	191,024	

C CNV-RNAseq2

	Reads	% Assigned reads
Total reads	367,985,967	----
Assigned	343,215,125	100%
-WT CNV RNA	342,918,966	99.9%
-*N. benthamiana*	294,522	0.09%
-Not annotated	1,637	0.0005%
Unassigned	24,770,842	

FIG 5 Summary of alignment of reads from CNVVLP-RNAseq, CNV-RNAseq1, and CNV-RNAseq2 with the *N. benthamiana* transcriptome. The numbers of reads from the CNVVLP-RNAseq (A), CNV-RNAseq1 (B), and CNV-RNAseq2 (C) data sets that aligned with the genes (CNV CP ORF and TBSV p19), the genome (WT CNV RNA), or the transcriptome (*N. benthamiana*) were determined using CLC Genomics Workbench (v6.5 and v7.5). Not annotated, no or little nucleotide sequence identity to other proteins could be detected during annotation.

encoded or mitochondrion-encoded RNAs. The most abundantly encapsidated nucleus-encoded mRNA (RPK = ~3,800) was an uncharacterized protein that showed strong sequence similarity to 7S signal recognition particle (SRP) RNA, and the most abundantly encapsidated mitochondrion-encoded RNA was 60S ribosomal protein L2 (RPK = ~9,800). Overall, chloroplast RNAs appeared to be much more efficiently encapsidated than nucleus-encoded or mitochondrion-encoded RNAs (Fig. 6A). Encapsidation of chloroplast RNA is consistent with the observation that CNV CP enters chloroplasts during agroinfiltration, as described above. The observation that mitochondrial RNAs are also encapsidated is likely a reflection of the entry of CNV CP into mitochondria, also described above. Taken together, the results indicate that CNV CP can assemble with a vast array of host RNA species in agroinfiltrated plants.

As will be described in more detail below, CNV CP also encapsidated the CP mRNA and TBSV p19 mRNA expressed during agroinfiltration. In fact, CNV CP was the most efficiently encapsidated of the mRNAs expressed in *N. benthamiana*, with an RPK value of ~43,000, i.e., 48,887 reads (Fig. 5A) per 1,140-nt coding region. TBSV p19 mRNA has an RPK value of ~24,000, i.e., 12,360 reads (Fig. 5A) per 518-nt coding region, making it more efficiently encapsidated than any nucleus- or mitochondrion-encoded mRNA. A measure of the efficiency of encapsidation of these RNAs relative to that of host RNAs is described further below.

Figure 6B summarizes the relative amounts of host RNAs encapsidated in CNV virions on the basis of the RPK values of individual transcript identifiers. Interestingly, some similarities with the CNVVLP-RNAseq data set (Fig. 6A) existed in the overall trend with respect to the levels of encapsidated host RNAs. Chloroplast rRNAs were again the most efficiently encapsidated, as were chloroplast mRNAs, in comparison to nucleus-encoded and mitochondrion-encoded rRNAs and mRNAs. It might be expected that chloroplast RNAs represent the most abundantly encapsidated RNAs in CNVVLPs since a large proportion of the CNV CP is present in chloroplasts (Fig. 4). However, chloroplast-located CNV CP represents only approximately 1 to 5% of the CP present in infected cells (38), but it is responsible for encapsidating the majority of host (chloroplast) RNAs. It is possible that assembly of CNV RNA occurs in the cytoplasm in a microenvironment that separates the CP from the host RNA species and, thus, viral RNA is preferentially encapsidated. CNV RNA may also contain sequences that facilitate the assembly process, which would contribute to the preferential encapsidation efficiency (see below). In addition, as has been suggested for the encapsidation of BMV RNA, the presence of the replicase could enhance the specificity of encapsidation (25, 48, 49), a process which, in the case of CNV, would not take place in chloroplasts, since CNV replicase does not enter chloroplasts; rather, it is targeted to peroxisomes (50).

As the total number of reads corresponding to encapsidated host RNAs was relatively low and the CNV preparation used for the analyses whose results are presented in Fig. 6B was not treated with RNase prior to deep sequencing analysis (although it was

A **CNVVLP-RNAseq**

Nuclear encoded RNA	RPK	#Reads	Chloroplast RNA	RPK	#Reads	Mitochondrial RNA	RPK	#Reads
rRNA			**rRNA**			**rRNA**		
26S	303,738	1,014,180	23S	761,100	2,137,931	26S	60,870	207,262
18S	51,935	91,822	16S	417,029	621,374	18S	44,098	83,831
mRNA			**mRNA**			**mRNA**		
Putative uncharacterized protein n=1 Tax=Ricinus communis RepID=B9S749_RICCO (probable) (7S SRP RNA)	3,839	1263	Ribulose bisphosphate carboxylase large chain (RbcL)	38,206	108,586	60S ribosomal protein L2, mitochondrial (probable)	9,795	4,241
Putative cyclin-D6-1 (probable)	1,831	3,807	YCF2_NICSY Protein ycf2 (probable)	33,558	42,056	Maturase-related protein n=29 Tax=core eudicotyledons RepID=Q5MA61_TOBAC (probable)	5,507	3,869
Ribulose bisphosphate carboxylase small chain 8B, chloroplastic (probable)	793	474	Protein ycf2 (probable)	28,720	28,651	Uncharacterized mitochondrial protein ymf17 (probable)	3,927	6,547
Ribulose bisphosphate carboxylase small chain 8B, chloroplastic (similar to)	771	389	DNA-directed RNA polymerase subunit beta' (probable)	13,122	11,963	NADH-ubiquinone oxidoreductase chain 4 (probable)	3,889	3,377
Gypsy/Ty-3 retroelement polyprotein (probable)	760	307	Maturase K (probable)	12,269	21,238	Ribosomal protein S14, mitochondrial (probable)	3,775	1,381
Heat shock cognate 70 kDa protein 2 (probable)	748	1,871	DNA-directed RNA polymerase subunit beta" (probable)	12,016	28,747	ATP synthase subunit alpha, mitochondrial	3,743	13,255
F-box protein SKIP14 (probable)	706	1,504	Photosystem II CP43 chlorophyll apoprotein	8,323	28,607	Ribosomal protein S13, mitochondrial (probable)	3,599	1,922
Retrovirus-related Pol polyprotein from transposon TNT 1-94 (probable)	672	1,048	RK2_NICTO 50S ribosomal protein L2, chloroplastic (probable)	6,501	10,697	Group II intron-encoded protein ltrA (probable)	3,572	2,729
DNA-damage-repair/toleration protein DRT100 (probable)	637	794	Photosystem I P700 chlorophyll a apoprotein A1	5,975	35,195	NADH-ubiquinone oxidoreduct-ase chain 1 n=3 Tax=core eudicotyledons RepID=Q5MA12 TOBAC (probable)	3,378	1,027
NF-X1-type zinc finger protein NFXL1 (probable)	637	2,315	NAD(P)H-quinone oxidoreductase subunit 2 B, chloroplastic (probable)	5,439	22,322	Ribosomal protein S19, mitochondrial (probable)	2,300	1,023

B **CNV-RNAseq1**

Nuclear encoded RNA	RPK	#Reads	Chloroplast RNA	RPK	#Reads	Mitochondrial RNA	RPK	#Reads
rRNA			**rRNA**			**rRNA**		
26S	301	1,005	23S	857.9	2,410	26S	42	143
18S	229.1	405	16S	867.1	1,292	18S	24.2	46
mRNA			**mRNA**			**mRNA**		
Putative uncharacterized protein n=2 Tax=*Vitis vinifera* RepID=A5B1A3_VITVI (probable)(retrotransposon)	32.7	44	Ribulose bisphosphate carboxylase large subunit (RbcL)	437.4	1,191	Cytochrome b-c1 complex subunit Rieske-2, mitochondrial (probable)	20.8	30
Peptidyl-prolyl cis-trans isomerase FKBP62 (PPI)	31.9	74	Photosystem II CP43 chlorophyll apoprotein	60.5	208	60S ribosomal protein L2	11.5	6
Putative uncharacterized protein n=4 Tax=*Vitis vinifera* RepID=A5BF11_VITVI (probable)(retrotransposon)	27.7	35	RR14_SOLBU 30S ribosomal protein S14, chloroplastic (probable)	60.5	39	Pentatricopeptide repeat-containing protein At4g01030, mitochondrial (probable)	9.8	29
Transposon Ty3-G Gag-Pol polyprotein (probable)	24.7	44	YCF2_NICSY Protein ycf2 (probable)	51.8	43	Maturase related protein n=29 Tax=core eudicotyledons RepID=Q5MA61_TOBAC	8.2	5
Putative uncharacterized protein n=1 Tax=*Vitis vinifera* RepID=A5AQF1_VITVI (probable)(retrotransposon)	24.7	11	Photosystem II CP47 chlorophyll apoprotein	43.5	94	Heat shock 70 kDa protein, mitochondrial (probable)	6.7	11
Putative uncharacterized protein n=1 Tax=*Vitis vinifera* RepID=A5BDH9_VITVI (probable)(retrotransposon)	24.0	13	Photosystem II DII (probable)	43.1	105	Ribosomal protein S13, mitochondrial (probable)	5.6	3
Retrotransposable element Tf2 155 kDa protein type 2 (probable)	19.9	22	Chlorophyll a-b binding protein 7, chloroplastic (probable)	38.0	31	Group II intron-encoded protein ltrA (probable)	5.5	4
Putative uncharacterized protein (Fragment) n=3 Tax=Vitis vinifera RepID=A5C2P7_VITVI (probable)	17.5	44	Photosystem I P700 chlorophyll a apoprotein A1	26.5	156	Threonine--tRNA ligase, mitochondrial (probable)	4.5	11
Nbv5tr6213785 (not annotated)	16.7	4	ATP synthase subunit beta	20.1	55	60S ribosomal protein L5, mitochondrial (probable)	4.2	4
Elongation factor 2	11.4	16	Photosystem I P700 chlorophyll a apoprotein A2	17.3	59	Cytochrome C oxidase subunit 1	4.0	9

Continued on following page

C

CNV-RNAseq2

Nuclear encoded RNA	RPK	#Reads	Chloroplast RNA	RPK	#Reads	Mitochondrial RNA	RPK	#Reads
rRNA			rRNA			rRNA		
26S	16,791	56,067	23s	29,045	81,617	26S	2,611	8,718
18S	29,241	51,143	16S	30,751	45,851	18S	2,546	4,840
mRNA			mRNA			mRNA		
Transposon Ty3-G Gag-Pol polyprotein (probable)	686	1860	Ribulose bisphosphate carboxylase large chain (RbcL)	22,437	56,662	Maturase related protein n=29 Tax=core eudicotyledons RepID=Q5MA61_TOBAC	361	254
Putative uncharacterized protein n=2 Tax=*Vitis vinifera* RepID=A5B1A3_VITVI (probable)(retrotransposon)	662	781	YCF2_NICSY Protein ycf2 (probable)	1,882.5	2,252	NADH-ubiquinone oxidoreductase chain 4 (probable)	305	276
Putative uncharacterized protein n=1 Tax=*Vitis vinifera* RepID=A5B4X8_VITVI (probable)(retrotransposon)	532	275	Photosystem I P700 chlorophyll a apoprotein A1	1,817	10,700	ATP synthase subunit alpha, mitochondrial	301	1066
Retrotransposable element Tf2 155 kDa protein type 2 (probable)	521	539	Photosystem II CP43 chlorophyll apoprotein	1,637	5,627	Ribosomal protein S13, mitochondrial (probable)	294	157
Putative uncharacterized protein n=4 Tax=*Vitis vinifera* RepID=A5BF11_VITVI (probable)(retrotransposon)	420	514	Photosystem II CP47 chlorophyll apoprotein	1,614	3,489	60S ribosomal protein L2	279	122
Putative uncharacterized protein n=1 Tax=*Vitis vinifera* RepID=A5BDH9_VITVI (probable)(retrotransposon)	402	233	Cytochrome b6 (probable)	1433	1386	Ribosomal protein S10, mitochondrial (probable)	241	185
Retrovirus-related Pol polyprotein from transposon 297 (probable)	280	401	Protein ycf2 (probable)	1,216.6	1,394	Orf107b protein n=3 Tax=Beta RepID=Q9MFC6_BETVU (probable)	203	123
RNA-directed DNA polymerase homolog (probable)	248	229	Photosystem I P700 chlorophyll a apoprotein A2	1,190	4,061	NADH-ubiquinone oxidoreductase chain 2 (probable)	183	216
DnaJ protein homolog (probable)	236	352	ATP synthase subunit beta	858	2,344	Group II intron-encoded protein ltrA (probable)	175	143
Hop-interacting protein THI031 n=1 Tax=Solanum lycopersicum RepID=G8Z260_SOLLC (probable)	232	178	Photosystem II DII (probable)	853	2,080	Ribosomal protein S19, mitochondrial (probable)	173	76

FIG 6 Major host RNAs encapsidated in pCNVCPpBin(+) VLPs and WT CNV particles. The numbers of reads of the indicated rRNAs or mRNAs aligned to sequences in the CNVVLP-RNAseq (A), CNV-RNAseq1 (B), and CNV-RNAseq2 (C) data sets are indicated. Alignments were performed against the *N. benthamiana* transcriptome. The CLC Genomics Workbench (v6.5) was used to perform the alignments using a length fraction of 0.5 and a similarity fraction of 0.95. All reads of less than 40 nt were discarded from the analysis. The number of reads per kilobase (RPK) value and the number of reads are given for each transcript, and the ordering is based on the RPK values. All *V. vinifera* sequences were determined by BLAST analysis to correspond to retrotransposon-like sequences, as outlined in Fig. S1 to S3 in the supplemental material. It is noted that several chloroplast and mitochondrial transcripts in the *N. benthamiana* transcriptome are multicistronic and are annotated according to the 5′-most gene. Thus, the number of reads for a given reference gene may include reads that mapped to other 3′ distally located genes in the reference transcript. Figures S1 and S3 in the supplemental material provide detailed information regarding the mapped sequences.

highly purified), we repeated the deep sequencing analysis using RNase-treated virions and Illumina sequencing and obtained more reads. In addition, the CNV preparation was from a different batch of infected plants. In this data set (CNV-RNAseq2) (Fig. 6C), approximately 296,000 host reads rather than the approximately 63,000 (including sequences not annotated) obtained in the CNV-RNAseq1 data set (Fig. 6B) were obtained, allowing a more statistically valid interpretation. The pattern of encapsidated host RNAs was remarkably similar, with chloroplast RNAs once again representing the majority by far. Moreover, 8 of the top 10 encapsidated chloroplast RNAs in the CNV-RNAseq2 data set were represented in the top 10 encapsidated chloroplast RNAs in the CNV-RNAseq1 data set. As for the encapsidated mitochondrial RNAs, only 4 of the top 10 encapsidated RNAs in CNV-RNAseq2 were the same as those in CNV-RNAseq1, possibly owing to the statistically significantly lower number of mitochondrial reads obtained in the CNV-RNAseq1 data set. Additionally, it is possible that some of the differences between the data sets may be due to the use of Illumina versus Ion Torrent sequence analysis. Also, the CNV-RNAseq1 data set may have included stable RNA species that copurified with CNV virions, since this data set was derived from CNV virions that were not treated with RNase.

It was also noted that many of the chloroplast and mitochondrial RNAs efficiently encapsidated in VLPs were also efficiently encapsidated by CNV during infection. For example, 7 of the top 10 encapsidated mitochondrial RNAs in the CNV-RNAseq2 data set were present in CNVVLP-RNAseq, and 5 of the top 10 encapsidated chloroplast RNAs in the CNVRNAseq2 data set were also present in CNVVLP-RNAseq (compare Fig. 6A and C). This may simply be a reflection of the abundances of these RNA species in pCNVCPpBin (+)-agroinfiltrated and CNV-infected plants, but a potential preference for encapsidating certain RNA species may also be present. Further research is required to ascertain these possibilities.

Retrotransposons are encapsidated by both pCNVCPpBin(+) VLPs and WT CNV particles. A most interesting finding in both CNV virion data sets was the prevalence of RNA sequences in

A

Retrotransposons encapsidated in pCNVCPpBin(+) VLPs

Retrotransposon	Type	Reads
Gypsy	LTR Class I	2059
Copia	LTR Class I	151
LINE-1	Non-LTR Class I	235
	Total	2445
	% Total assigned reads (n=5,239,655)	0.047%

B

Retrotransposons encapsidated in WT CNV particles

Retrotransposon	Type	Reads
Gypsy	LTR Class I	217
Copia	LTR Class I	35
LINE-1	Non-LTR Class I	2
	Total	254
	% Total assigned host RNA reads (n=63,509)	0.4%

C

Retrotransposons encapsidated in WT CNV particles

Retrotransposon	Type	Reads
Gypsy	LTR Class I	3,223
Copia	LTR Class I	429
LINE-1	Non-LTR Class I	41
	Total	3,693
	% Total assigned host RNA reads (n=296,159)	1.3%

FIG 7 Retrotransposons encapsidated in pCNVCPpBin(+) VLPs and in WT CNV particles. The numbers of reads from the CNVVLP-RNAseq (A), CNV-RNAseq1 (B), and CNV-RNAseq2 (C) data sets that aligned with retrotransposons are shown. Retrotransposons were identified on the basis of their annotations using the search terms Copia, Gypsy, TY3, Tf2, TNT, retrotransposable element, retrotransposon, retroelement, LINE-1, and transposon and then grouped into the corresponding superfamilies (i.e., Gypsy, Copia, and LINE-1). It is noted that several annotated transcripts that were not identified as transposons in the *N. benthamiana* transcriptome were found to have sequence similarity to known transposons, but reads that mapped to these are not indicated in the data sets. Thus, the number of reads should be regarded as the smallest number of reads corresponding to retrotransposons. The *n* values for the percentages of total assigned reads and total assigned host RNA reads are from Fig. 5.

nucleus-encoded encapsidated RNAs that correspond to retrotransposons or retrotransposon-like RNAs (Fig. 6B and C). The retrotransposons included both major superfamilies of long terminal repeat (LTR) class I transposons, Gypsy and Copia, which are abundant in plant genomes, as well as LINE-1 transposons, which are non-LTR class I transposons also known to be present in plants (51–55) (Fig. 7). In CNV-RNAseq2, these elements represented the top 8 encapsidated nucleus-encoded RNAs. In CNV-RNAseq1, they represented 6 of the top 10. Retrotransposons or retrotransposon-like sequences were not as abundant in CNV VLPs, where they represented only 2 of the top 10 encapsidated RNAs. Figure 7 summarizes the retrotransposons encapsidated in CNV VLPs and CNV virions. In CNV VLPs (Fig. 7A), retrotransposons represented approximately 0.047% of the total assigned reads, with LTR class I Gypsy family-related transposons being the most abundant. In CNV virions, retrotransposon sequences represented approximately 0.4% and 1.3% of the total assigned host reads in the CNV-RNAseq1 and CNV-RNAseq2 data sets, respectively. In both these cases, LTR class I Gypsy elements were the most abundant, as in CNV VLPs. It is noted that during the analysis of retrotransposons in CNV virions, it became apparent that the *N. benthamiana* genome likely encodes more retrotransposons than are annotated. This was found when analyzing the top 10 encapsidated nuclear host RNAs in CNV virions. For example, in a pBLAST search of the translated sequences of 4 of the top 10 encapsidated RNAs which were uncharacterized sequences matching *Vitis vinifera* RNAs (see Fig. 6C, column 1), the presence of conserved sequences corresponding to the various proteins found in retrotransposons was observed. For example, as described in Fig. S3 in the supplemental material, a putative uncharacterized protein labeled $n = 2$, Tax = *Vitis vinifera*, RepID = A5B1A3_VITVI contained signature domains for the Gag polyprotein and pepsin/retropepsin superfamilies. Similarly, a putative uncharacterized protein labeled $n = 1$, Tax = *Vitis vinifera*, RepID = A5BDH9_VITVI contained a nucleotide sequence with similarity to that of the *Beta vulgaris* TY3/Gypsy retrotransposon (GenBank accession number JX455094.1) as well as showed domain features of the pepsin/retropepsin superfamily. Figure 7 provides data only for reads mapped to reference genes with annotations referring to retrotransposons, so these are likely an underestimate of the contribution of retrotransposons to encapsidated RNAs in CNV virions.

The observation that CNV VLPs contain an approximately 10- to 28-fold lower level of retrotransposons than CNV virions could suggest that retrotransposons become more prominent during CNV infection. It has been reported that stress and pathogen infection, including virus infection, can induce retrotransposons (56–59), and should this occur during CNV infection, the increased availability of retrotransposons may result in a greater proportional encapsidation.

In FHV, retrotransposons were found to account for 0.1% of total encapsidated RNA (23). As was suggested for FHV, the presence of retrotransposon sequences in CNV particles could possibly provide a means for horizontal transmission of retrotransposons should CNV infect a new host. Horizontal gene transfer is well-known to occur among organisms, including plants, and is believed to catalyze genome evolution. Diao et al. (60) documented the horizontal transmission of a DNA transposon, and more recently, widespread horizontal transmission of retrotransposons has been inferred to occur in multiple plant species (61). Pathogens, including viruses, were suggested to play a role in horizontal transmission, and our data would be consistent with this possibility.

Confirmation of encapsidation of host RNAs by VLPs and CNV particles. RT-PCR was conducted on selected encapsidated CNV and VLP RNAs to confirm encapsidation and as a means to assess the integrity of the RNA, as it is possible that only small

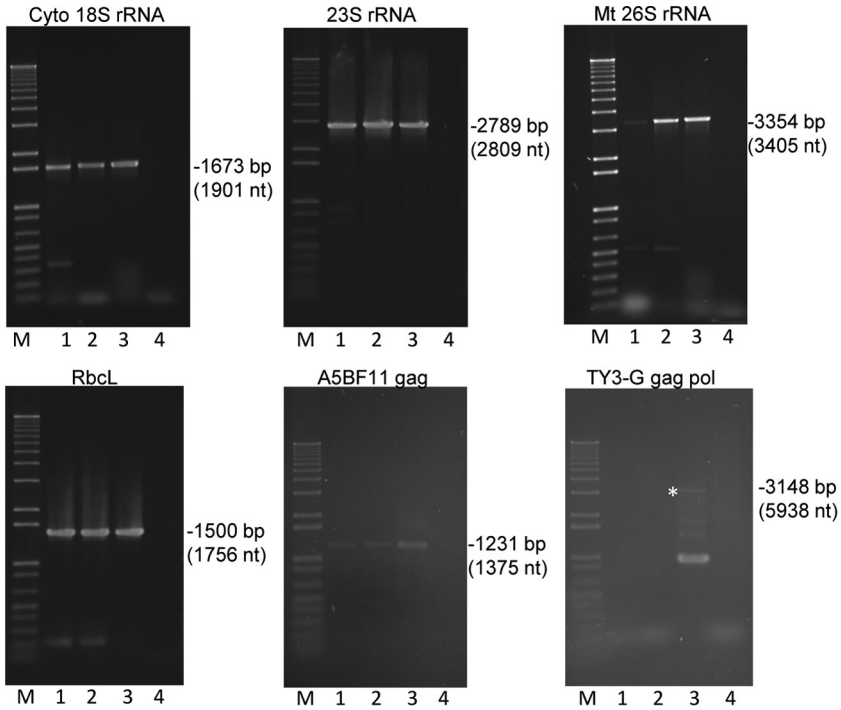

FIG 8 RT-PCR analysis of several pCNVCPpBin(+)- and CNV-encapsidated host RNAs. RT-PCR followed by agarose gel electrophoresis was conducted on each of the indicated cytoplasmic (Cyto), chloroplastic, or mitochondrial (Mt) RNA species detected in pCNVCPpBin(+)- and CNV-encapsidated host RNAs by NGS analysis to confirm the presence and assess the integrity of the RNA in RNase-treated VLPs or CNV particles (see Fig. 6 for the corresponding genes and abbreviations). Lanes 1, total leaf RNA used as a positive control; lanes 2, pCNVCPpBin(+) VLP RNA; lanes 3, WT CNV virion RNA; lanes 4; water used in place of RNA as a negative PCR control; lanes M, molecular size markers. The sizes of the RT-PCR products are shown to the right of each gel, and the number in parentheses below the size of the RT-PCR products indicates the size of the full-length transcript or that of the largest transcript recorded in the *N. benthamiana* database. PCR primers are described in Table S1 in the supplemental material, and the transcript identifiers for each RNA are provided in Fig. S1 in the supplemental material. The asterisk in the TY3-G Gag Pol gel highlights the 3,148-bp fragment for which primers were synthesized. This fragment was sequenced and demonstrated to correspond to this gene. The GenBank accession numbers used to prepare the primers are as follows: AJ236016 for cytoplasmic 18S rRNA, Z00044 for 23S rRNA, and BA000042 for mitochondrial 26S rRNA. Transcript identifiers from the *N. benthamiana* Sydney database are as follows: Nbvtr56230855 for RbcL, Nbv5tr6274273 for A5BF11 Gag, and Nbvtrv56274272 for Ty3-G Gag Pol.

fragments of host RNA are encapsidated, as suggested on the basis of the short read lengths obtained during standard deep sequencing analysis. Figure 8 shows that long cDNAs for cytoplasmic 18S rRNA, chloroplast 23S rRNA, RbcL, mitochondrial 26S rRNA, and two reference genes from Fig. 6 containing retrotransposon-like features (A5BF11 and TY3-G Gag Pol) could be amplified. The sizes of the cDNAs ranged from approximately 1,200 nt to ~3,300 nt, indicating that large portions of the RNAs are present in VLPs and CNV virions. In the case of 18S rRNA, 23S rRNA, RbcL, and mitochondrial 26S rRNA, the RT-PCR products correspond to nearly full-length RNA. The CNV genome is ~4.7 kb, and this would be expected to be the size for optimal encapsidation. Previous work has shown that up to 5.4 kb of RNA can be encapsidated by CNV (62) and that T=3 particles can encapsidate RNAs as small as 2.1 kb (36). Thus, encapsidation of full-length cytoplasmic 18S rRNA could likely be in T=3 particles produced by the cytoplasmic 41-kDa protein. However, encapsidation of 23S rRNA and RbcL mRNA by cleaved chloroplast CP was not expected since the fully cleaved stromally located protein is expected to be capable of producing only T=1 particles, as deduced from previous studies of CNV mutants lacking the R domain (35–37). The volume of a T=1 particle is expected to be able to accommodate RNA, which would be approximately 1 kb or less. Thus, it is possible that the nearly full-length 23S rRNA and RbcL mRNA could correspond to RNA species that were encapsidated by full-length CP following chloroplast breakage during infection or artifactually during the virus isolation procedure. The droplet digital PCR (ddPCR) analyses described below utilized primer sets producing PCR products of only ~100 to 400 nt in length, so these data are likely a more valid representation of the contribution of host RNAs to virion RNA populations (Fig. 9).

CNV CP mRNA and TBSV p19 mRNA are efficiently encapsidated in pCNVCPpBin(+)-agroinfiltrated plants. Viral RNAs are believed to contain sequences and/or structures that are preferentially recognized during virus infection to ensure efficient encapsidation of the genome and to reduce the possibility of encapsidation of host RNA, as described above. CNV CP mRNA and TBSV p19 mRNA are encapsidated in pCNVCPpBin(+) VLPs, on the basis of the results of deep sequencing analysis (Fig. 5A), and represent 2 of the 3 most abundantly encapsidated mRNAs. As CNV or the closely related TBSV p19 mRNAs may contain sequences that promote assembly, we wished to investigate the encapsidation efficiency of CP mRNA with respect to that of TBSV p19 mRNA and specific encapsidated host RNAs in pCNVCPpBin(+) VLPs. *N. benthamiana* plants were coagroinfiltrated with pCNVCPpBin(+) and TBSV p19, and at 5 dpi leaves were ground to a fine powder in liquid nitrogen. A portion of the material was used to extract total leaf RNA, and the remaining portion was used

A REE of specific host RNAs with respect to CNV CP mRNA in pCNVCPpBin(+) VLPs

Encapsidated RNA	Ratio of accumulation of the indicated RNA in total leaf RNA with respect to CNV CP mRNA		Ratio of encapsidation of the indicated RNA in virion RNA with respect to CNV CP mRNA		REE of the indicated RNA with respect to CNV CP mRNA (%)		Average REE
	Exp.1	Exp.2	Exp.1	Exp.2	Exp.1	Exp.2	
TBSV p19	4.5	2.5	2.4	1.9	53.3%	76.0%	64.6%
Nuclear encoded							
18S rRNA	48.7	8.6	0.2	0.19	0.41%	2.2%	1.3%
Chloroplastic							
23S rRNA	22	21.3	1.25	4.7	5.7%	22.1%	13.9%
RbcL	26.5	53	1.4	3.35	5.3%	6.2%	5.7%
Mitochondrial							
26S rRNA	22.5	24.6	0.27	0.96	1.2%	3.9%	2.6%

B REE of specific host RNAs with respect to CNV RNA in CNV virions

Encapsidated RNA	Ratio of accumulation of the indicated RNA in total leaf RNA with respect to WT CNV RNA		Ratio of encapsidation of the indicated RNA in virion RNA with respect to WT CNV RNA		REE of the indicated RNA with respect to WT CNV RNA (%)		Average REE
Nuclear encoded	Rep .1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2	
18S rRNA	0.005	0.005	$3 X10^{-6}$	$7 X 10^{-7}$	0.06%	0.014%	0.037%
PPI	0.0018	0.004	$5 X 10^{-7}$	$1 X10^{-6}$	0.04%	0.02%	0.03%
Chloroplastic							
23S rRNA	0.014	0.017	$4 X10^{-6}$	$2 X10^{-6}$	0.03%	0.01%	0.02%
RbcL	0.043	0.05	$2 X10^{-4}$	$8 X10^{-5}$	0.46%	0.16%	0.31%
Mitochondrial							
26S rRNA	0.0032	0.0034	$3 X10^{-6}$	$2 X10^{-6}$	0.09%	0.06%	0.075%

FIG 9 REEs of specific host RNAs with respect to those of CNV CP mRNA in CNVCPpBin(+) VLPs or CNV RNA in WT CNV virions. ddPCR was conducted to determine the absolute numbers of the indicated RNA species in total leaf RNA or virions from pCNVCPpBin(+)-agroinfiltrated leaves (A) or from CNV-infected leaves (B). The primer sets used for ddPCR are described in Table S1 in the supplemental material. The REE was calculated as described in the text. (A) Each experiment (Exp.) was conducted 2 times (i.e., the data were obtained from independent inoculations) and included 2 technical repeats (i.e., independent dilutions from the same reverse transcriptase reaction) per experiment, which were averaged and are presented here. (B) Data were obtained from a single CNV inoculation. Rep. 1 and Rep. 2, separate reverse transcription reactions from the same virion and total leaf RNA samples. The numbers are averages for independent dilutions from the same reverse transcription reaction. PPI, peptidyl prolyl *cis-trans* isomerase, which was found to be encapsidated in CNV particles (Fig. 6B).

for VLP purification. Purified VLPs were treated with RNase, and then VLP RNA was extracted. Both VLP RNA and total leaf RNA extractions utilized DNase I to remove any contaminating DNA. ddPCR using gene-specific primers (see Table S1 in the supplemental material) was used to quantitate the number of CNV CP, TBSV p19, or specific encapsidated host transcripts (Fig. 9A) present in equivalent amounts of total leaf RNA and VLP RNA. The following formula, previously used to estimate the relative encapsidation efficiency (REE) of viral RNAs (10), was used to determine the REE of CNV CP mRNA compared to that of cytoplasmic TBSV p19 and 18S rRNA, chloroplastic 23S rRNA and the RbcL transcript, and mitochondrial 26S rRNA as follows: the ratio of the number of molecules of encapsidated host RNA of interest relative to that of CNV CP RNA divided by the ratio of the number of molecules of encapsidated RNA species of interest in total leaf RNA relative to that of CNV CP RNA. It was found that 18S rRNA, 23S rRNA, RbcL RNA, and 26S rRNA were encapsidated, on average, at only 1.3%, 13.9%, 5.7%, and 2.6% of the efficiency of CNV CP mRNA, respectively (Fig. 9A), suggesting that the CNV CP mRNA contains a site (or sites) that is relatively more efficiently recognized by CNV CP for VLP assembly and, moreover, that such a site or sites may serve as one or more of the nucleation sites for assembly of CNV RNA during infection. This is a significant finding, since little is known about the CNV RNA sequences involved in the specificity of viral RNA assembly. Interestingly, the TBSV p19 mRNA was encapsidated at approximately 65% of the efficiency of CNV CP mRNA. The nucleotide sequence of the TBSV p19 ORF is highly similar (86% identical) to that of its

homolog, CNV p20. The efficient encapsidation of TBSV p19 mRNA thus suggests that the CNV p20 ORF may also contain one or more nucleation sites for CNV genomic RNA assembly.

Using the same method described above for encapsidation in VLPs, we also examined the REE of specific host RNAs with respect to that of CNV RNA in CNV virions. Figure 9B shows that each of the host RNAs examined (i.e., 18S rRNA, peptidyl prolyl *cis-trans* isomerase RNA, 23S rRNA, RbcL RNA, and 26S rRNA) is encapsidated, on average, at a very low level in comparison to that of CNV RNA, with relative encapsidation efficiencies ranging from 0.02% for 23S rRNA to 0.31% for RbcL. This is consistent with the low level of these RNAs in CNV virion RNA preparations and further suggests that CNV genomic RNA contains sequences that promote assembly specificity.

The notion that viral RNAs contain specific sequences and/or structures that promote encapsidation is supported by our observation that CNV CP and TBSV p19 mRNAs are encapsidated relatively more efficiently than other RNAs transcribed in CNVCPpBin(+)-infiltrated plants. However, a variety of other nonviral RNAs, including the major cellular rRNAs and numerous host mRNAs, are also encapsidated, suggesting that encapsidation can occur, albeit less efficiently, in the absence of these specific viral RNA sequences. Nevertheless, the CNV virions produced during infection predominantly contain viral RNA. It is interesting that the majority of host RNAs encapsidated in CNV virions are chloroplast encoded. Nonspecific encapsidation of chloroplast RNAs may occur in the absence of a functional replicase, as has been previously suggested for BMV (25, 48, 49), whereas more specific encapsidation may occur in its presence in the cytoplasm. Also, CNV is replicated within spherules derived from peroxisomes in the cells of *N. benthamiana* (50). Should encapsidation take place in the vicinity of replication, the predominant RNA is likely to be viral RNA, and thus, virions would contain a low level of host RNA species. It is likely that a number of factors contribute to the specificity of CNV encapsidation, including the presence of a high-affinity site(s) for CP binding (in the CP- and CNV p20-coding regions) and the relative predominance of viral RNA at the site of replication. Our data support both of these hypotheses.

Encapsidation of sequences derived from retrotransposons or retrotransposons that are likely amplified during infection suggests that viral infection may promote genome evolution and also contribute to the horizontal transmission of retrotransposons to other naive hosts, thus influencing the evolution of these hosts as well.

ACKNOWLEDGMENTS

This work was partially supported by an NSERC Discovery Grant (10R82367) as well as an NSERC Supplementary Accelerator Grant (10R68048).

We thank Basudev Ghoshal, Hala Khalil, and Paul Wiersma for kindly providing help with the bioinformatics analyses.

REFERENCES

1. **Annamalai P, Rao AL.** 2007. In vivo packaging of brome mosaic virus RNA3, but not RNAs 1 and 2, is dependent on a *cis*-acting 3′ tRNA-like structure. J Virol **81:**173–181. http://dx.doi.org/10.1128/JVI.01500-06.
2. **Basnayake VR, Sit TL, Lommel SA.** 2009. The red clover necrotic mosaic virus origin of assembly is delimited to the RNA-2 trans-activator. Virology **384:**169–178. http://dx.doi.org/10.1016/j.virol.2008.11.005.
3. **Bink HH, Schirawski J, Haenni AL, Pleij CW.** 2003. The 5′-proximal hairpin of turnip yellow mosaic virus RNA: its role in translation and encapsidation. J Virol **77:**7452–7458. http://dx.doi.org/10.1128/JVI.77.13.7452-7458.2003.
4. **Borodavka A, Tuma R, Stockley PG.** 2012. Evidence that viral RNAs have evolved for efficient, two-stage packaging. Proc Natl Acad Sci U S A **109:**15769–15774. http://dx.doi.org/10.1073/pnas.1204357109.
5. **Choi YG, Rao AL.** 2003. Packaging of brome mosaic virus RNA3 is mediated through a bipartite signal. J Virol **77:**9750–9757. http://dx.doi.org/10.1128/JVI.77.18.9750-9757.2003.
6. **Damayanti TA, Tsukaguchi S, Mise K, Okuno T.** 2003. *cis*-Acting elements required for efficient packaging of brome mosaic virus RNA3 in barley protoplasts. J Virol **77:**9979–9986. http://dx.doi.org/10.1128/JVI.77.18.9979-9986.2003.
7. **Frolova E, Frolov I, Schlesinger S.** 1997. Packaging signals in alphaviruses. J Virol **71:**248–258.
8. **Hemmer O, Dunoyer P, Richards K, Fritsch C.** 2003. Mapping of viral RNA sequences required for assembly of peanut clump virus particles. J Gen Virol **84:**2585–2594. http://dx.doi.org/10.1099/vir.0.19247-0.
9. **Johnson SF, Telesnitsky A.** 2010. Retroviral RNA dimerization and packaging: the what, how, when, where, and why. PLoS Pathog **6:**e1001007. http://dx.doi.org/10.1371/journal.ppat.1001007.
10. **Kemler I, Barraza R, Poeschla EM.** 2002. Mapping the encapsidation determinants of feline immunodeficiency virus. J Virol **76:**11889–11903. http://dx.doi.org/10.1128/JVI.76.23.11889-11903.2002.
11. **Kwon SJ, Park MR, Kim KW, Plante CA, Hemenway CL, Kim KH.** 2005. *cis*-Acting sequences required for coat protein binding and in vitro assembly of potato virus X. Virology **334:**83–97. http://dx.doi.org/10.1016/j.virol.2005.01.018.
12. **McBride MS, Schwartz MD, Panganiban AT.** 1997. Efficient encapsidation of human immunodeficiency virus type 1 vectors and further characterization of *cis* elements required for encapsidation. J Virol **71:**4544–4554.
13. **Murakami S, Terasaki K, Narayanan K, Makino S.** 2012. Roles of the coding and noncoding regions of Rift Valley fever virus RNA genome segments in viral RNA packaging. J Virol **86:**4034–4039. http://dx.doi.org/10.1128/JVI.06700-11.
14. **Qu F, Morris TJ.** 1997. Encapsidation of turnip crinkle virus is defined by a specific packaging signal and RNA size. J Virol **71:**1428–1435.
15. **Rochon D, Siegel A.** 1984. Chloroplast DNA transcripts are encapsidated by tobacco mosaic virus coat protein. Proc Natl Acad Sci U S A **81:**1719–1723. http://dx.doi.org/10.1073/pnas.81.6.1719.
16. **Zhong W, Dasgupta R, Rueckert R.** 1992. Evidence that the packaging signal for nodaviral RNA2 is a bulged stem-loop. Proc Natl Acad Sci U S A **89:**11146–11150. http://dx.doi.org/10.1073/pnas.89.23.11146.
17. **Zimmern D, Butler PJ.** 1977. The isolation of tobacco mosaic virus RNA fragments containing the origin for viral assembly. Cell **11:**455–462. http://dx.doi.org/10.1016/0092-8674(77)90064-2.
18. **Gopal R, Venter PA, Schneemann A.** 2014. Differential segregation of nodaviral coat protein and RNA into progeny virions during mixed infection with FHV and NoV. Virology **454-455:**280–290. http://dx.doi.org/10.1016/j.virol.2014.03.003.
19. **Patel N, Dykeman EC, Coutts RH, Lomonossoff GP, Rowlands DJ, Phillips SE, Ranson N, Twarock R, Tuma R, Stockley PG.** 2015. Revealing the density of encoded functions in a viral RNA. Proc Natl Acad Sci U S A **112:**2227–2232. http://dx.doi.org/10.1073/pnas.1420812112.
20. **Liu Y, Wang C, Mueller S, Paul AV, Wimmer E, Jiang P.** 2010. Direct interaction between two viral proteins, the nonstructural protein 2C and the capsid protein VP3, is required for enterovirus morphogenesis. PLoS Pathog **6:**e1001066. http://dx.doi.org/10.1371/journal.ppat.1001066.
21. **Pickett GG, Peabody DS.** 1993. Encapsidation of heterologous RNAs by bacteriophage MS2 coat protein. Nucleic Acids Res **21:**4621–4626. http://dx.doi.org/10.1093/nar/21.19.4621.
22. **Johnson KN, Tang L, Johnson JE, Ball LA.** 2004. Heterologous RNA encapsidated in Pariacoto virus-like particles forms a dodecahedral cage similar to genomic RNA in wild-type virions. J Virol **78:**11371–11378. http://dx.doi.org/10.1128/JVI.78.20.11371-11378.2004.
23. **Routh A, Domitrovic T, Johnson JE.** 2012. Host RNAs, including transposons, are encapsidated by a eukaryotic single-stranded RNA virus. Proc Natl Acad Sci U S A **109:**1907–1912. http://dx.doi.org/10.1073/pnas.1116168109.
24. **Suluja E, Strokowskaja L, Zagorski-Ostoja W, Palucha A.** 2005. Virus-like particles of potato leafroll virus as potential carrier system for nucleic acids. Acta Biochim Pol **52:**699–702.

25. **Annamalai P, Rao AL.** 2005. Replication-independent expression of genome components and capsid protein of brome mosaic virus in planta: a functional role for viral replicase in RNA packaging. Virology **338:**96–111. http://dx.doi.org/10.1016/j.virol.2005.05.013.
26. **Siegel A.** 1971. Pseudovirions of tobacco mosaic virus. Virology **46:**50–59. http://dx.doi.org/10.1016/0042-6822(71)90005-5.
27. **Rochon D, Lommel S, Martelli GP, Rubino L, Russo M.** 2012. Tombusviridae, p 1111–1138. *In* King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (ed), Virus taxonomy. Ninth report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, San Diego, CA.
28. **Rochon DM, Tremaine JH.** 1988. Cucumber necrosis virus is a member of the tombusvirus group. J Gen Virol **69:**395–400. http://dx.doi.org/10.1099/0022-1317-69-2-395.
29. **Rochon DM, Tremaine JH.** 1989. Complete nucleotide sequence of the cucumber necrosis virus genome. Virology **169:**251–259. http://dx.doi.org/10.1016/0042-6822(89)90150-5.
30. **Rochon DM, Johnston JC.** 1991. Infectious transcripts from cloned cucumber necrosis virus cDNA: evidence for a bifunctional subgenomic mRNA. Virology **181:**656–665. http://dx.doi.org/10.1016/0042-6822(91)90899-M.
31. **Johnston JC, Rochon DM.** 1990. Translation of cucumber necrosis virus RNA in vitro. J Gen Virol **71**(Pt 10)**:**2233–2241. http://dx.doi.org/10.1099/0022-1317-71-10-2233.
32. **Sit TL, Johnston JC, ter Borg MG, Frison E, McLean MA, Rochon D.** 1995. Mutational analysis of the cucumber necrosis virus coat protein gene. Virology **206:**38–48. http://dx.doi.org/10.1016/S0042-6822(95)80017-4.
33. **Johnston JC, Rochon DM.** 1996. Both codon context and leader length contribute to efficient expression of two overlapping open reading frames of a cucumber necrosis virus bifunctional subgenomic mRNA. Virology **221:**232–239. http://dx.doi.org/10.1006/viro.1996.0370.
34. **Li M, Kakani K, Katpally U, Johnson S, Rochon D, Smith TJ.** 2013. Atomic structure of cucumber necrosis virus and the role of the capsid in vector transmission. J Virol **87:**12166–12175. http://dx.doi.org/10.1128/JVI.01965-13.
35. **Katpally U, Kakani K, Reade R, Dryden K, Rochon D, Smith TJ.** 2007. Structures of T=1 and T=3 particles of cucumber necrosis virus: evidence of internal scaffolding. J Mol Biol **365:**502–512. http://dx.doi.org/10.1016/j.jmb.2006.09.060.
36. **Kakani K, Reade R, Katpally U, Smith T, Rochon D.** 2008. Induction of particle polymorphism by cucumber necrosis virus coat protein mutants in vivo. J Virol **82:**1547–1557. http://dx.doi.org/10.1128/JVI.01976-07.
37. **Reade R, Kakani K, Rochon D.** 2010. A highly basic KGKKGK sequence in the RNA-binding domain of the cucumber necrosis virus coat protein is associated with encapsidation of full-length CNV RNA during infection. Virology **403:**181–188. http://dx.doi.org/10.1016/j.virol.2010.03.045.
38. **Xiang Y, Kakani K, Reade R, Hui E, Rochon D.** 2006. A 38-amino-acid sequence encompassing the arm domain of the cucumber necrosis virus coat protein functions as a chloroplast transit peptide in infected plants. J Virol **80:**7952–7964. http://dx.doi.org/10.1128/JVI.00153-06.
39. **Hui E, Xiang Y, Rochon D.** 2010. Distinct regions at the N-terminus of the cucumber necrosis virus coat protein target chloroplasts and mitochondria. Virus Res **153:**8–19. http://dx.doi.org/10.1016/j.virusres.2010.06.021.
40. **Ghoshal K, Theilmann J, Reade R, Sanfacon H, Rochon D.** 2014. The cucumber leaf spot virus p25 auxiliary replicase protein binds and modifies the endoplasmic reticulum via N-terminal transmembrane domains. Virology **468-470:**36–46. http://dx.doi.org/10.1016/j.virol.2014.07.020.
41. **Voinnet O, Rivas S, Mestre P, Baulcombe D.** 2003. An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J **33:**949–956. http://dx.doi.org/10.1046/j.1365-313X.2003.01676.x.
42. **Kakani K, Sgro JY, Rochon D.** 2001. Identification of specific cucumber necrosis virus coat protein amino acids affecting fungus transmission and zoospore attachment. J Virol **75:**5576–5583. http://dx.doi.org/10.1128/JVI.75.12.5576-5583.2001.
43. **Robbins MA, Reade RD, Rochon DM.** 1997. A cucumber necrosis virus variant deficient in fungal transmissibility contains an altered coat protein shell domain. Virology **234:**138–146. http://dx.doi.org/10.1006/viro.1997.8635.
44. **Hillman BI, Hearne P, Rochon D, Morris TJ.** 1989. Organization of tomato bushy stunt virus genome: characterization of the coat protein gene and the 3′ terminus. Virology **169:**42–50. http://dx.doi.org/10.1016/0042-6822(89)90039-1.
45. **Yukawa M, Tsudzuki T, Sugiura M.** 2005. The 2005 version of the chloroplast DNA sequence from tobacco (*Nicotiana tabacum*). Plant Mol Biol Rep **23:**359–365. http://dx.doi.org/10.1007/BF02788884.
46. **Sugiyama Y, Watase Y, Nagase M, Makita N, Yagura S, Hirai A, Sugiura M.** 2005. The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. Mol Genet Genomics **272:**603–615. http://dx.doi.org/10.1007/s00438-004-1075-8.
47. **Nakasugi K, Crowhurst RN, Bally J, Wood CC, Hellens RP, Waterhouse PM.** 2013. De novo transcriptome sequence assembly and analysis of RNA silencing genes of Nicotiana benthamiana. PLoS One **8:**e59534. http://dx.doi.org/10.1371/journal.pone.0059534.
48. **Chaturvedi S, Rao AL.** 2014. Live cell imaging of interactions between replicase and capsid protein of brome mosaic virus using bimolecular fluorescence complementation: implications for replication and genome packaging. Virology **464-465:**67–75. http://dx.doi.org/10.1016/j.virol.2014.06.030.
49. **Rao AL, Chaturvedi S, Garmann RF.** 2014. Integration of replication and assembly of infectious virions in plant RNA viruses. Curr Opin Virol **9:**61–66. http://dx.doi.org/10.1016/j.coviro.2014.09.008.
50. **Rochon D, Singh B, Reade R, Theilmann J, Ghoshal K, Alam SB, Maghodia A.** 2014. The p33 auxiliary replicase protein of cucumber necrosis virus targets peroxisomes and infection induces de novo peroxisome formation from the endoplasmic reticulum. Virology **452-453:**133–142. http://dx.doi.org/10.1016/j.virol.2013.12.035.
51. **Domingues DS, Cruz GM, Metcalfe CJ, Nogueira FT, Vicentini R, Alves CDS, Van Sluys MA.** 2012. Analysis of plant LTR-retrotransposons at the fine-scale family level reveals individual molecular patterns. BMC Genomics **13:**137. http://dx.doi.org/10.1186/1471-2164-13-137.
52. **Feschotte C, Jiang N, Wessler SR.** 2002. Plant transposable elements: where genetics meets genomics. Nat Rev Genet **3:**329–341. http://dx.doi.org/10.1038/nrg793.
53. **Suoniemi A, Tanskanen J, Schulman AH.** 1998. Gypsy-like retrotransposons are widespread in the plant kingdom. Plant J **13:**699–705. http://dx.doi.org/10.1046/j.1365-313X.1998.00071.x.
54. **Ungerer MC, Strakosh SC, Stimpson KM.** 2009. Proliferation of Ty3/gypsy-like retrotransposons in hybrid sunflower taxa inferred from phylogenetic data. BMC Biol **7:**40. http://dx.doi.org/10.1186/1741-7007-7-40.
55. **Voytas DF, Cummings MP, Koniczny A, Ausubel FM, Rodermel SR.** 1992. Copia-like retrotransposons are ubiquitous among plants. Proc Natl Acad Sci U S A **89:**7124–7128. http://dx.doi.org/10.1073/pnas.89.15.7124.
56. **Beguiristain T, Grandbastien MA, Puigdomenech P, Casacuberta JM.** 2001. Three Tnt1 subfamilies show different stress-associated patterns of expression in tobacco. Consequences for retrotransposon control and evolution in plants. Plant Physiol **127:**212–221.
57. **Casacuberta JM, Vernhettes S, Audeon C, Grandbastien MA.** 1997. Quasispecies in retrotransposons: a role for sequence variability in Tnt1 evolution. Genetica **100:**109–117. http://dx.doi.org/10.1023/A:1018309007841.
58. **Jones RB, Song H, Xu Y, Garrison KE, Buzdin AA, Anwar N, Hunter DV, Mujib S, Mihajlovic V, Martin E, Lee E, Kuciak M, Raposo RA, Bozorgzad A, Meiklejohn DA, Ndhlovu LC, Nixon DF, Ostrowski MA.** 2013. LINE-1 retrotransposable element DNA accumulates in HIV-1-infected cells. J Virol **87:**13307–13320. http://dx.doi.org/10.1128/JVI.02257-13.
59. **Takeda S, Sugimoto K, Otsuki H, Hirochika H.** 1998. Transcriptional activation of the tobacco retrotransposon Tto1 by wounding and methyl jasmonate. Plant Mol Biol **36:**365–376. http://dx.doi.org/10.1023/A:1005911413528.
60. **Diao X, Freeling M, Lisch D.** 2006. Horizontal transfer of a plant transposon. PLoS Biol **4:**e5. http://dx.doi.org/10.1371/journal.pbio.0040005.
61. **El Baidouri M, Carpentier MC, Cooke R, Gao D, Lasserre E, Llauro C, Mirouze M, Picault N, Jackson SA, Panaud O.** 2014. Widespread and frequent horizontal transfers of transposable elements in plants. Genome Res **24:**831–838. http://dx.doi.org/10.1101/gr.164400.113.
62. **Reade R, Delroux K, Macdonald K, Sit TL, Lommel SA, Rochon D.** 2001. Spontaneous deletion enhances movement of a cucumber necrosis virus based chimera expressing the red clover necrotic mosaic virus movement protein gene. Mol Plant Pathol **2:**13–25. http://dx.doi.org/10.1046/j.1364-3703.2001.00045.x.